BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images may include (but are not limited to):

- BLACK BORDERS
- TEXT CUT OFF AT TOP, BOTTOM OR SIDES
- FADED TEXT
- ILLEGIBLE TEXT
- SKEWED/SLANTED IMAGES
- COLORED PHOTOS
- BLACK OR VERY BLACK AND WHITE DARK PHOTOS
- GRAY SCALE DOCUMENTS

IMAGES ARE BEST AVAILABLE COPY.

As rescanning documents *will not* correct images, please do not report the images to the Image Problem Mailbox.

Proc. Natl. Acad. Sci. USA
Vol. 81, pp. 3998–4002, July 1984
Biochemistry

Use of peptide synthesis to probe viral antigens for epitopes to a resolution of a single amino acid

(antigenic determinant/foot-and-mouth disease virus)

H. MARIO GEYSEN*, ROB H. MELOEN†, AND SIMON J. BARTELING†

*Commonwealth Serum Laboratories, 45 Poplar Road, Parkville, Melbourne, Australia 3052; and †Central Veterinary Institute, 39 Houtribweg, Lelystad, The Netherlands

Communicated by G. J. V. Nossal, March 12, 1984

ABSTRACT **A procedure is described for rapid concurrent synthesis on solid supports of hundreds of peptides, of sufficient purity to react in an enzyme-linked immunosorbent assay. Interaction of synthesized peptides with antibodies is then easily detected without removing them from the support. In this manner an immunogenic epitope of the immunologically important coat protein of foot-and-mouth disease virus (type O_1) is located with a resolution of seven amino acids, corresponding to amino acids 146–152 of that protein. Then, a complete replacement set of peptides in which all 20 amino acids were substituted in turn at every position within the epitope was synthesized, and the particular amino acids conferring specificity for the reaction with antibody were determined. It was found that the leucine residues at positions 148 and 151 were essential for reaction with antisera raised against intact virus. A lesser contribution was derived from the glutamine and alanine residues at positions 149 and 152, respectively. Aside from the practical significance for locating and examining epitopes at high resolution, these findings may lead to better understanding of the basis of antigen–antibody interaction and antibody specificity.**

Recombinant DNA technology now makes possible by deduction from the determined nucleotide sequences reliable amino acid sequences of biologically important proteins. However, methods for identifying the loci in a protein that constitute the antigenic and immunogenic epitopes are few and time consuming and form the bottleneck to further rapid progress. Immunogenic epitopes are defined as those parts of a protein that elicit the antibody response when the whole protein is the immunogen. These immunogenic epitopes are believed to be confined to a few loci on the molecule (1–3). On the other hand, a region of a protein molecule to which an antibody can bind is defined as an antigenic epitope. Antisera prepared against chemically synthesized peptides corresponding to short linear tracts of the total polypeptide sequence have been shown to react well with the native protein (4–9). However, interactions were also found to occur even when the site of interaction did not correspond to an immunogenic epitope of the native protein. This has been interpreted to mean that the number of immunogenic epitopes of a protein is less than the number of antigenic epitopes (4). Conversely, since antibodies produced against the native protein are, by definition, directed to the immunogenic epitopes, it follows that peptides reacting with these antibodies must contain elements of the epitopes. From a study of the few proteins for which the determinants have been accurately mapped, it is postulated that a determinant may consist of a single element (continuous) or of more than one element brought together from linearly distant regions of the polypeptide chain by the folding of that chain as it exists in the native state (discontinuous) (10). Systematic mapping of all the detectable reactive elements of a protein by the chemical synthesis of overlapping segments has until now been severely limited by the scale of the synthetic and testing capability required (10, 11). Smith and co-workers (12, 13) circumvented the decoupling and purification steps by combining solid-phase peptide synthesis and solid-phase radioimmunoassay using the same solid support.

We describe here the concurrent synthesis of all 208 possible overlapping hexapeptides covering the total 213-amino acid sequence of the immunologically important coat protein (VP1) of foot-and-mouth disease virus (FMDV), type O_1 (Fig. 1). The peptides, still attached to the support used for their synthesis, were tested for antigenicity by an ELISA using a variety of antisera. After identification of a hexapeptide reactive with antibody raised against the intact virus, all 120 hexapeptides representing the complete single point amino acid replacement set were synthesized and tested for retention of antigenicity. By this method a whole virus epitope was examined at a resolution of a single amino acid.

MATERIALS AND METHODS

Synthesis of Peptides. Polyethylene rods (diameter, 4 mm; length, 40 mm) immersed in a 6% (vol/vol) aqueous solution of acrylic acid were γ irradiated at a dose of 1,000,000 rads (1 rad = 0.01 gray) (15). Rods so prepared were assembled into a polyethylene holder with the format and spacing of a microtiter plate. Subsequent reactions at the tips of the rods were carried out in a Teflon tray with a matrix of wells to match the rod spacing. Conventional methods of solid-phase peptide chemistry (16, 17) were used to couple N^{α}-*t*-butyloxycarbonyl-L-lysine methyl ester to the polyethylene/polyacrylic acid via the N^{ϵ}-amino group of the side chain. Carboxy substitution of the support was determined by treating NH_2-lysine(OMe)-polyethylene/polyacrylic acid with ^{14}C-labeled butyric acid and was found to be 0.15–0.2 nmol/mm^2. Removal of the *t*-butyloxycarbonyl group was followed by the coupling of *t*-butyloxycarbonyl-L-alanine to complete a peptide-like spacer. Successive amino acids were added as dictated by the sequence to be synthesized. At the completion of the final coupling reaction, and after removal of the *t*-butyloxycarbonyl protecting group, the terminal amino group was acetylated with acetic anhydride in dimethylformamide/triethylamine. All *N,N*-dicyclohexylcarbodiimide-mediated coupling reactions were carried out in dimethylformamide in the presence of *N*-hydroxybenzotriazole. The following side-chain protecting groups were used: *O*-benzyl for threonine, serine, aspartic acid, glutamic acid, and tyrosine; carbobenzoxy for lysine; tosyl for arginine; 4-methylbenzyl for cysteine; and 1-benzyloxycarbonylamido-2,2,2-trifluoroethyl for histidine. Side-chain-protecting groups were removed by treatment with borontris(trifluor-

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

Abbreviations: FMDV, foot-and-mouth disease virus; P_i/NaCl, phosphate-buffered saline.

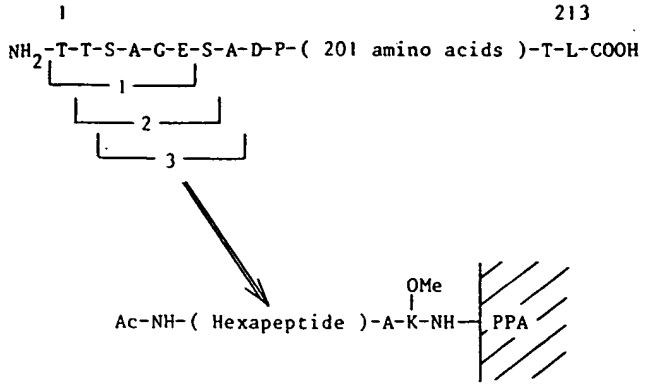

FIG. 1. The 213-amino acid sequence of VP1 (FMDV, type O_1) as translated by Kurz *et al.* (14) was subdivided into hexapeptide units, and each was synthesized on a separate polyethylene support in the orientation, and with a dipeptide spacer, as shown. Peptides are numbered according to the position of the NH-terminal amino acid within the VP1 sequence. PPA, polyethylene/polyacrylic acid.

acetate) in trifluoroacetic acid for 90 min at room temperature (18). After hydrolysis with HCl/propionic acid, sequences included in the synthesis as controls were analyzed to confirm that, although coupling at each stage had occurred, it was incomplete for several of the amino acids, notably arginine. Before testing by ELISA, support-coupled peptides were washed several times with phosphate-buffered saline (P_i/NaCl).

Antisera. Antisera against the intact virus particle were prepared by immunizing rabbits with 50 μg of inactivated, density gradient-purified virus in complete Freund's adjuvant. The animals were bled 3–4 weeks after the single inoculation. Anti-virus-subunit serum was prepared by inoculating rabbits three times, 3–4 weeks apart, with 10 μg of acid-disrupted purified virus, initially in complete Freund's and subsequently in incomplete Freund's adjuvant. The polypeptide VP1 was separated from the mixture of proteins obtained from urea-disrupted purified virus by isoelectric focusing (19). It was eluted from the gel with 8 M urea and dialyzed against P_i/NaCl, and antiserum was raised in rabbits as described for the virus subunit. Antiserum for scan 3 (see Fig. 2) was that used for scan 2 after absorption with purified virus (1500 μg of whole virus was incubated with 1 ml of serum for 72 hr at 4°C), and all virus-bound antibodies were removed by centrifugation.

ELISA. Support-coupled peptides were precoated with 10% horse serum/10% ovalbumin/1% Tween 80 in P_i/NaCl for 1 hr at 37°C to block nonspecific absorption of antibodies. Overnight incubation at 4°C in a 1:40 dilution of antiserum in the preincubation mixture was followed by three washes in 0.05% Tween 80/P_i/NaCl. Reaction for 1 hr at 37°C with a 1:50,000 dilution of goat anti-rabbit IgG coupled to horseradish peroxidase in the preincubation mixture was again followed by extensive washing with P_i/NaCl/Tween to remove excess conjugate. The presence of antibody was detected by reaction for 45 min with a freshly prepared developing solution (40 mg of *o*-phenylenediamine and 20 μl of hydrogen peroxide in 100 ml of phosphate buffer, pH 5.0), and the color produced was read in a Titertek Multiskan (Flow Laboratories, Melbourne, Australia) at 420 nm. Prior to retesting, bound antibody was removed from the peptides by washing peptides three times at 37°C in 8 M urea/0.1% 2-mercaptoethanol/0.1% sodium dodecyl sulfate and then several times with P_i/NaCl.

RESULTS

Identification of a Virus Particle-Associated Immunogenic Epitope. All 208 possible hexapeptides from the amino acid sequence of the VP1 protein of FMDV type O_1 were synthesized in duplicate. The amino acid sequence had been deduced from the nucleotide sequence of the VP1 gene (14). The results obtained for all the synthesized hexapeptides when tested by ELISA with six different antisera are shown in Fig. 2. Antisera used in the test were as follows: two different anti-(intact virus, type O_1), a virus-absorbed anti-(intact virus, type O_1), an anti-(virus subunit, type O_1), an anti-(isolated virus protein VP1, type O_1), and, as a control, an anti-(intact virus, type C_3). The two anti-intact virus sera tested, scans 1 and 2, show the extremes in the reactivity patterns found. Large quantitative differences in the individual animal responses to an identical antigen preparation have been reported before, but these scans highlight the variability possible in the antibody composition between sera. Examination of scans 1, 2, and 3 shows that antibodies reactive with hexapeptide numbers 146 and 147 are present in anti-intact particle sera (scans 1 and 2) but completely absent after absorption of the sera with purified virus (scan 3). Presumably, scan 3 registers those antibodies raised against epitopes expressed in denatured virions that are not present on the surface of the intact virion. Activities to hexapeptides 146 and 147 were not observed in the anti-subunit serum (scan 4) and were only weakly present in the anti-VP1 serum (scan 5). That some activity was found in the anti-VP1 serum

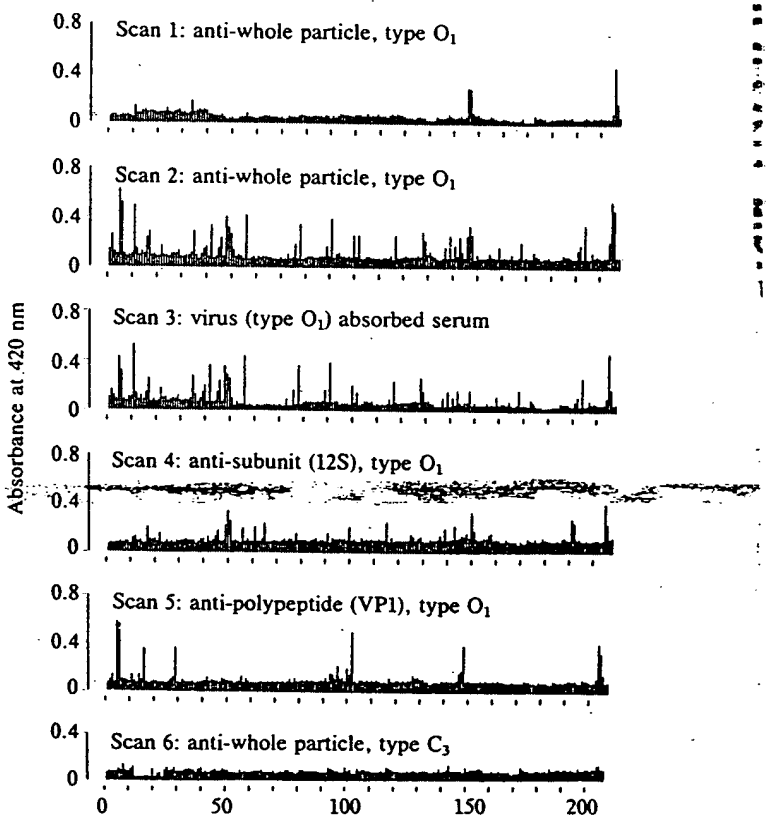

FIG. 2. Antigenic profiles (scans). Results are shown as vertical lines proportional to the extinction obtained in the antibody-binding ELISA test, plotted above the number giving the location within the VP1 sequence of the NH_2-terminal amino acid of each peptide. Antisera used to produce the scans shown were as follows: 1 and 2, two different anti-whole virus particle, type O_1; 3, anti-whole virus particle (as used in 2) after absorption with purified intact virus; 4, anti-virus subunit, type O_1; 5, anti-VP1, type O_1; 6, anti-whole virus particle, type C_3. It should be noted that, because the sequence of VP1 contains 20 alanine residues, 20 of the peptides synthesized match for seven amino acids. However, the frequency of reactive peptides from this group was not significantly different from the overall frequency (0.2 compared with 0.16) and therefore not considered further.

possibly accounts for the immunizing capacity, albeit weak, of the isolated protein (20). It should be noted however that another anti-VP1 serum tested, while retaining a strong activity at position 148, showed no activity at positions 146 and 147. Comparison of scan 3 with scan 2 (absorbed compared with nonabsorbed) shows that, in addition to the loss of activity to peptides 146 and 147, some reduction in activity to peptides 5, 6, and 206 also occurred. Of these, activity to 5 and 6 was not found in all the anti-intact virus sera tested, but activity to 206 was invariably present. From this we conclude that of the peptides found to be reactive, the pair at 146 (G-D-L-Q-V-L) and 147 (D-L-Q-V-L-A) [in this paper, amino acids are identified by the single-letter code (21)] constitute or are part of the principal immunogenic epitope, with the element at 206 (V-A-P-V-K-P) contributing to a lesser epitope. This is consistent with the observations of others (5, 22). Scan 6 shows the absence of reactivity in an antiserum produced against a different serotype of the virus.

Extending the Resolution of the Epitope at Peptides 146/147 to a Single Amino Acid. From the preceding data, we were unable to distinguish between two possibilities: (*i*) the epitope is contained in the five amino acids common to peptides 146 and 147—i.e., D-L-Q-V-L—or (*ii*) the epitope is represented by the "sum" of the two hexapeptides—i.e., G-D-L-Q-V-L-A. To extend the resolution, all 120 possible hexapeptides differing from peptide 146 (G-D-L-Q-V-L) by only a single amino acid were synthesized. Each of the other 19 common amino acids was substituted in each of the six amino acid positions within the peptide. Positions at which all or at least the majority of substitutions result in a loss of antibody-binding activity indicate those residues that are important for the specificity and binding to antibody. The ELISA activity obtained for each of the 120 peptides when serum 48 (anti-intact virus particle) was used in the test are shown in Fig. 3. The relative activities (with respect to the parent sequence) determined for each peptide for two different anti-intact virus sera, nos. 31 and 48, are given in Table 1. To determine the contribution of the alanine residue (carboxyl terminus of peptide 147) toward reactivity and/or specificity, a further 20 peptides were synthesized. Each of these peptides consisted of the complete sequence of 146 (G-D-L-Q-V-L) with one of the 19 possible amino acids added to the carboxyl terminus and synthesized as described before. When serum 31 was used in the test, activity was retained for seven of the amino acids. Relative values expressed in the same way as given in Table 1 were as follows: A (parent amino acid), 99; D, 55; E, 36; G, 45; N, 95; Q, 98; S, 44. With serum 48, activity was retained for four amino acids: A (parent amino acid), 94; G, 30; S, 47; T, 39.

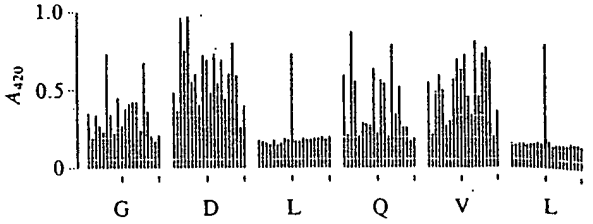

FIG. 3. Antibody-binding activity. The result for each peptide is shown as a vertical line proportional to the ELISA extinction obtained. Every group of 20 lines corresponds to the complete replacement set for one of the six amino acid positions in the hexapeptide G-D-L-Q-V-L. Within each group of 20 lines, the left-hand line corresponds to the substitution of the original residue by alanine (A), and the successive lines are then in alphabetic order according to the single letter code for the amino acids.

DISCUSSION

Interpretation of Data. In choosing to adopt the procedure for peptide synthesis as described, we made several assumptions.

1. To detect antibodies, the quantity of peptide of a defined sequence need only be in the pmol range (5). Assuming a worst-case overall yield of 1% for an eight-step synthesis (two linking and six sequence amino acids), an initial level of 1 nmol of reactive group per support would satisfy the above condition.

2. High purity for the peptide used in the detection of antibodies is not a necessary condition. The majority of serological tests rely on the specificity of antibodies to detect a given antigen in the presence of large amounts of irrelevant protein.

3. Except for cases in which either all or none of the peptides react, a large number of the peptides would effectively act as negative controls in the test. With adjacent peptides sharing a common sequence of five amino acids, the observation of peaks above a generally uniform background level would indicate a valid test.

4. Many of the antibodies elicited by immunization with an intact virus result from presentation of epitopes in fully or partially denatured form. Such antibodies may bind to synthetic peptides *in vitro* but not to the virion itself. They are therefore assumed to be less relevant to virus neutralization

Table 1. Relative antibody-binding activities of peptides derived from the parent sequence G-D-L-Q-V-L

Serum	Parent residue	Activity when substituted with amino acid																			
		A	C	D	E	F	G	H	I	K	L	M	N	P	Q	R	S	T	V	W	Y
31	G	29					**90**	14	27	12		32	34	41	29		50				
	D	22	21	143	95	110	28	65		12	65	15	58	10	69		38	62			
	L										**79**										
	Q			64	14								13		80						
	V	62		33	52				26		29	59			45		49	43	89		
	L										**119**										
48	G	11					**88**	10		32		18	24	25	26		77	14			
	D	37	12	136	92	137	52	62	21	87	81	37	89	49	80	29	63	104	60		21
	L										**88**										
	Q	60		117	52					68		53	49		102	10	45				
	V	52		40	63	42			56	82	68	88	34		106	33	91	98	81		14
	L										**105**										

Antibody-binding activities are shown for all peptides that gave an extinction significantly above background. Values for each peptide are expressed as a percentage of the mean activity of the six parent sequences synthesized as a part of each replacement set. Values given boldface type correspond to those obtained for the parent sequence. No activity was detected when the antiserum used was prepared against the heterologous FMDV type.

than are antibodies that bind to virions as well as to peptides.

The extinction obtained in an ELISA for a given peptide depends on the concentration of the antibody population with the correct specificity for reaction. It is essentially independent of the peptide density expressed as reacting groups per mm^2 of support (unpublished data). The difference in the extinction obtained with peptides synthesized with densities varying over two orders of magnitude is similar to the 10–30% variation observed between replicate synthesis (unpublished data). The extinction may also be expected to depend greatly on the affinity between peptide and reacting antibody, but this remains to be verified, although the overnight reaction would tend to minimize differences. Antigenic profiles of the FMDV VP1 (Fig. 2) were interpreted to define an antigenic peptide as one giving an ELISA extinction significantly above the background level of the test. On the other hand, in the testing of replacement nets (Fig. 3), the concentration of the reactive antibody population is constant and effectively of one specificity. Therefore, the extinctions observed are interpreted to reflect the mean affinity of the reacting antibody population for the peptide.

Immunogenicity of FMDV Virus Protein (VP1). Scan 5 of Fig. 2 identifies six immunogenic regions defined in terms of epitopes on the isolated protein eliciting antibodies capable of binding to the corresponding synthetic peptide. Scan 4 shows that, for the same protein as a part of the virus subunit, additional regions (principally 50–70 and 191–197) are immunogenic. Scan 2 shows that, during the course of the immune response to whole virus, most of the protein can be immunogenic. In contrast, scan 1 shows a response to only a very limited number of epitopes. What has become clear from these and other results (unpublished) is that different animals do not necessarily respond to all of the epitopes on a given antigen. In addition, the immunogenic response of an individual animal will be complicated if the antigen is readily broken down as is known to happen to FMDV (23, 24). The animal is exposed not only to the intact virus but also to subunits and possibly even to the isolated viral proteins. Each of these different states could present different epitopes to the immune system. Epitopes can be identified with a particular state of the antigen by testing the peptides with antisera specific to that state.

An Immunogenic Epitope at High Resolution. Antibodies raised against a particular immunogenic epitope will have a combining site (paratope) complementary to the structure of that epitope. An antibody population directed to the same epitope (allowing for variation in the expression of antibodies by the immune reponse) will have common features in the combining sites essential for binding to that epitope. A peptide that, in one of its many conformations in thermal equilibrium *in vitro*, has a structure sufficiently similar to the form of the epitope against which antibody was raised *in vivo* will bind to the antibody. Modification of a reacting peptide by amino acid substitution will define the limits for interaction with antibody. By so "mapping" the antibody-combining site, it is possible to infer properties of the antigen to which this antibody population is complementary. Using polyclonal antisera, it was not expected that a rigorous requirement for particular amino acids in particular positions would be observed. It is clear that, whatever the diversity of the antibodies involved in the interaction, the requirement for a given amino acid in certain positions is absolute for most or all of the antibodies present. It is also clear that the specificity range found for the two different antisera is remarkably similar, differing mainly in the hierarchy of preference for amino acids at the nonessential position. As judged from the limitation to replacements at some position within the sequence G-D-L-Q-V-L-A, the whole-virus epitope may be considered to be X-X-L-Q-X-L-A, where X is nonessential, letters in boldface type indicate an absolute requirement, and letters in lightface type indicate a contributing amino acid.

These findings suggest a different interpretation of the characteristics of epitopes. The antigenic specificity of the epitope represented by amino acids 146–152 within the VP1 protein of FMDV is largely dependent on the leucine residues at positions 148 and 151. These are hydrophobic residues and would not normally be expected to protrude from the protein surface. This suggests the possibility that the immune system responds to a local protein conformation that is different from that expected to represent the global energy minimum. The energy for antigen–antibody binding may be derived from the positive entropy term associated with the transfer of hydrophobic residues from a hydrophilic (aqueous) environment to within the antibody-combining site.

Scope of the Described Approach to Epitope Mapping. Although our results have been presented for a single protein only, the agreement with results of others in locating a viral epitope within the region encompassing amino acids 141–160 of VP1 is excellent (5, 22). The further resolution obtained by Rowlands *et al.* (25) from the comparison of the sequences of the VP1s of three antigenic variants of a single virus type (A_{12}) showed that amino acid substitution at positions 148 and/or 153 would affect the ability to react with specific antibody. This result is in good agreement with our results for subtype O_1, where positions 148 and 151 were critical to the immunogenicity of the epitope. We expect that the systematic approach as outlined, when applied to a broader spectrum of proteins, will contribute greatly to our understanding of the nature of epitopes and their interaction with the immune system.

We thank Mr. Jan Briaire for his enthusiastic and skilled technical assistance with the synthesis of the peptides, Mr. Jan Meyer for assistance with the ELISA, and Dr. Dick Voskamp of the Technical University, Delft, for valuable advice on aspects of the peptide chemistry. This work is the result of a collaborative project between the Commonwealth Serum laboratories and the Central Veterinary Institute and was conducted at the latter institute.

1. Crumpton, M. J. (1974) in *The Antigens*, ed. Sela, M. (Academic, New York), Vol. 2, pp. 1–78.
2. Benjamini, E., Michaeli, D. & Young, J. D. (1972) *Curr. Top. Microbiol. Immunol.* **58**, 85–134.
3. Atassi, M. Z. (1975) *Immunochemistry* **12**, 423–438.
4. Green, N., Alexander, H., Olson, O., Alexander, S., Shinnick, T. M., Sutcliffe, J. G. & Lerner, R. A. (1982) *Cell* **28**, 477–487.
5. Bittle, J. L., Houghten, R. A., Alexander, H., Shinnick, T. M., Sutcliffe, J. G., Lerner, R. A., Rowlands, D. J. & Brown, F. (1982) *Nature (London)* **298**, 30–33.
6. Dreesman, G. R., Sanchez, Y., Ionescu-Matiu, I., Sparrow, J. T., Six, H. R., Peterson, D. L., Hollinger, F. B. & Melnick, J. L. (1982) *Nature (London)* **295**, 158–160.
7. Prince, A. M., Ikram, H. & Hopp, T. P. (1982) *Proc. Natl. Acad. Sci. USA* **79**, 579–582.
8. Lerner, R. A., Green, N., Alexander, H., Liu, F., Sutcliffe, J. G. & Shinnick, T. M. (1981) *Proc. Natl. Acad. Sci. USA* **78**, 3403–3407.
9. Neurath, A. R., Kent, S. B. H. & Strick, N. (1982) *Proc. Natl. Acad. Sci. USA* **79**, 7871–7875.
10. Atassi, M. Z. (1978) *Immunochemistry* **15**, 909–936.
11. Kazim, A. L. & Atassi, M. Z. (1980) *Biochem. J.* **191**, 261–264.
12. Smith, J. A., Hurrell, J. G. R. & Leach, S. J. (1977) *Immunochemistry* **14**, 565–568.
13. Hurrell, J. G. R., Smith, J. A. & Leach, S. J. (1978) *Immunochemistry* **15**, 297–302.
14. Kurz, C., Forss, S., Kupper, H., Strohmaier, K. & Schaller, H. (1981) *Nucleic Acids Res.* **9**, 1919–1931.
15. Muller-Schulte, D. & Horster, F. A. (1982) *Polym. Bull.* **7**, 77–81.
16. Erickson, B. W. & Merrifield, R. B. (1976) in *The Proteins*, eds. Neurath, H. & Hill, R. L. (Academic, New York), Vol. 2, pp. 255–527.

17. Meienhofer, J. (1973) in *Hormonal Proteins and Peptides*, ed. Li, C. H. (Academic, New York), Vol. 2, pp. 45–267.
18. Pless, J. & Bauer, W. (1973) *Angew. Chem.* **85**, 142.
19. Barteling, S. J., Wagenaar, F. & Gielkens, A. L. J. (1982) *J. Gen. Virol.* **62**, 357–361.
20. Kleid, D. G., Yansura, D., Small, B., Dowbenko, D., Moore, D. M., Grubman, M. J., McKercher, P. D., Morgan, D. O., Robertson, B. H. & Bachrach, H. L. (1981) *Science* **214**, 1125–1129.
21. IUPAC-IUB Commission on Biochemical Nomenclature (1968) *Eur. J. Biochem.* **5**, 151–153.
22. Pfaff, E., Mussgay, M., Bohm, H. O., Schulz, G. E. & Schaller, H. (1982) *EMBO J.* **1**, 869–874.
23. Brown, F. (1981) *Trends Biochem. Sci.* **6**, 325–327.
24. Meloen, R. H. & Briaire, J. (1980) *J. Gen. Virol.* **55**, 107–116.
25. Rowlands, D. J., Clarke, B. E., Carroll, A. R., Brown, F., Nicholson, B. H., Bittle, J. L., Houghten, R. A. & Lerner, R. A. (1983) *Nature (London)* **306**, 694–697.

Selection of Immunogenic Peptides for Antisera Production

In order to produce antisera reactive with proteins, a peptide is selected from a translated cDNA or protein sequence and is then synthesized, purified (at least partially), and conjugated to a carrier protein. Strategies for selecting immunogenic peptides and carrier proteins as well as various means of coupling peptides to carriers are described. The actual coupling protocols, a typical immunization protocol, and a procedure for preparing a peptide affinity column for antibody purification are included in *UNIT 9.4*. A flowchart for the preparation of antipeptide antibodies (antibodies reactive with a synthetic peptide) is shown in Figure 9.3.1.

SELECTION OF AN IMMUNOGENIC PEPTIDE

To prepare antibodies against synthetic peptides, a peptide must first be selected. This is the most critical step in obtaining an antibody that reacts with the native antigen. In practice, a 10- to 15-residue peptide sequence inferred from a cDNA sequence or from an N-terminal amino acid sequence is selected. If possible, sequences should be avoided that are likely to be identical or highly homologous to those in the animal to be immunized (usually rabbits). After synthesis and purification, the peptide is cross-linked to a carrier protein such as keyhole limpet hemocyanin (KLH).

If a peptide sequence to be utilized for antisera production is not from the terminal regions of the protein, selection is based on predicting antigenic sites. It is presumed that the sites accessible to reactivity with antipeptide antibodies are exposed on the surface of the protein; these sites are likely to be more common in flexible regions of the protein (Westhof et al., 1984), and are more likely to be found on reverse turns or loop structures (Dyson et al., 1985). Computer algorithms

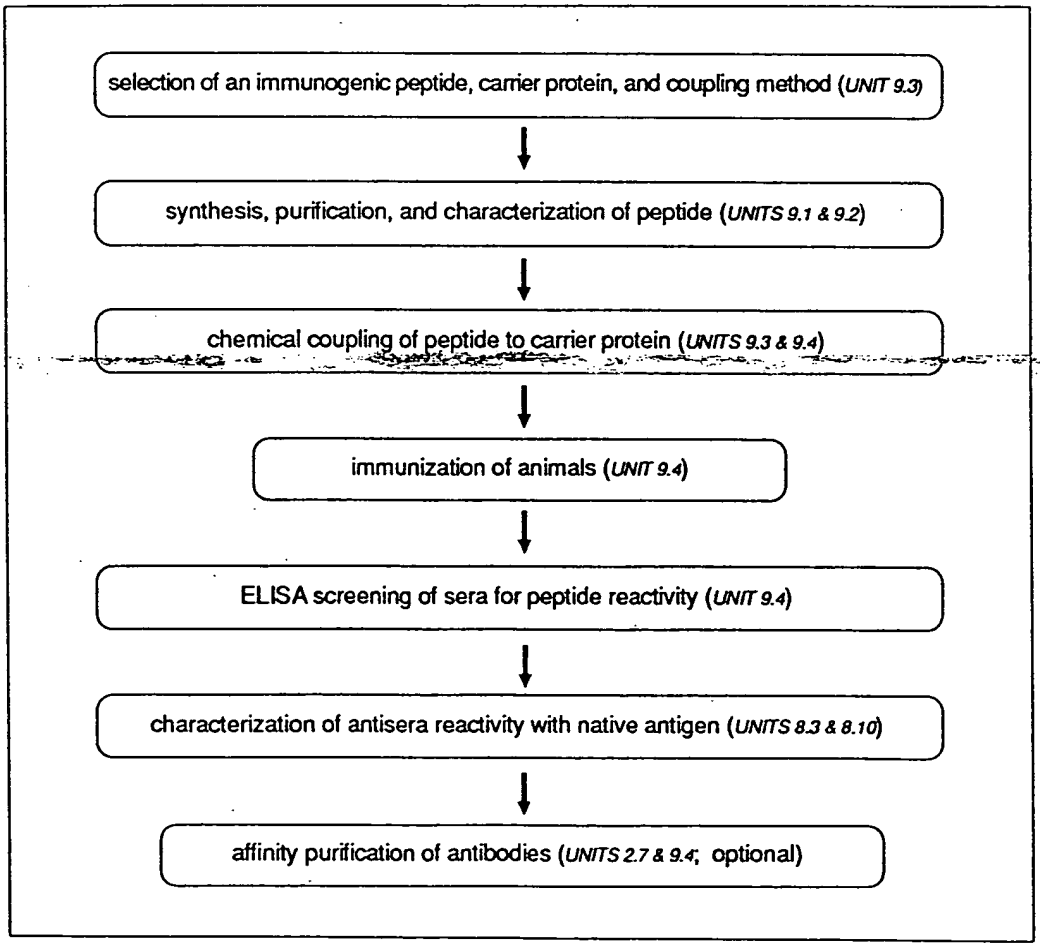

Figure 9.3.1 Flow chart for preparation and analysis of antipeptide antibodies.

estimate hydrophilicity (Kyte and Doolittle, 1982; Hopp and Woods, 1983), flexibility (Karplus and Schultz, 1985), and secondary structure of the protein along its polypeptide chain; they can help identify surfaces on the protein that have appropriate characteristics for antigenic sites. Antisera produced against peptides derived from such regions are more likely to react with the native protein than peptides derived from regions without these characteristics.

Selection of a C-Terminal Peptide

Because the C termini of proteins are often more mobile than the rest of the molecule and are frequently exposed on the protein surface, this region is usually accessible to antibody-combining sites. This is particularly true for detergent-solubilized transmembrane proteins with C-terminal cytoplasmic tails—i.e., MHC class I and class II molecules, and T cell receptor molecules. This type of peptide can be coupled to the carrier in a straightforward manner using *m*-maleimidobenzoyl-*N*-hydroxysuccinimide ester (MBS) via a Cys residue that has been added to the N terminus of the selected peptide. By coupling the peptide via its N-terminal end to the carrier protein, the peptide will be exposed in a fashion similar to that found in the native antigen. Obviously, if any other Cys residues are present, they will also couple; therefore, an alternate coupling procedure may be more appropriate (see coupling methods).

Unfortunately there will be instances when the C-terminal sequence will not be the preferable immunogen—e.g., if its sequence is part of a transmembrane region and is thus too hydrophobic. This can be determined by examining a hydropathicity plot of the sequence (Kyte and Doolittle, 1982).

Selection of an N-Terminal Peptide

As with the C terminus, the N terminus is often exposed in the native protein; therefore, peptides selected from this region are often useful for making antibodies reactive with the intact protein. In this case, if MBS coupling is used, the Cys residue should be present on the C terminus. By coupling the peptide via the C-terminal end to the carrier protein, it will be oriented as it is in the native protein. Unfortunately, the N terminus of a protein may have a post-translational modification, such as acetylation or fatty acid acylation. Moreover, if the protein sequences are derived from cDNA sequences, the leader sequence must not be confused with the authentic N terminus. It is usually possible to locate potential leader sequences using an algorithm derived by von Heijne (1986). If it is known that the authentic N terminus is acetylated, the peptide can then be acetylated during synthesis to reproduce the structure in the native protein.

Selection of an Internal Peptide Sequence

Selection of a peptide from an internal part of the protein sequence can be aided by the use of algorithms to predict those regions most likely to be exposed on the surface of the protein. The only information needed is the primary amino acid sequence. Two of the methods discussed below are based on the fact that antibody-reactive sites are usually located in externally exposed, hydrophilic regions of proteins. A third method relies on secondary structure predictions.

Using the first algorithmic method, a hydrophilicity value is assigned for each overlapping six-amino-acid segment of the protein sequence based on the average of the hydrophilicity values (Table 9.3.1) of the amino acids in that segment. The highest point of average local hydrophilicity is usually located in or near an antigenic determinant. A computer program written in Basic is available for analysis (Hopp and Woods, 1983).

An alternative algorithm evaluates the hydrophobic and hydrophilic tendencies of a polypeptide chain based on water vapor free-energy transfers and the interior versus exterior distributions of amino acid side chains. Values for each amino acid are listed in the second column of Table 9.3.1. A computer program for calculating this hydropathicity profile has been written by Kyte and Doolittle (1982). This profile is useful for determining exterior and interior regions of a native protein, as well as for locating signal sequences and transmembrane sequences.

A third algorithm is an empirical method that relies on a library of known structures to determine the frequency with which each amino acid occurs in the various conformational states (i.e., α-helix, β-sheet, β-turn, or all other structural forms; Chou and Fasman, 1974). Using these frequencies, predictions can be made about secondary structure for a given sequence. For making antipeptide sera, regions that are predicted to form turns or loops, or extended sequences (20 to 25 residues) that have a very high probability for formation of an α-helix, are useful. A computer program for performing this analysis can be found in

Corrigan and Huang (1982).

Although no data exists to prove this point, it would seem wise to avoid choosing peptides containing predicted polysaccharide attachment sites, most notably the sequence Asn-X-Ser or Asn-X-Thr, which predict the presence of an Asn-linked polysaccharide. It is likely that the presence of polysaccharide moieties at such sites in the native protein would interfere with antibody accessibility.

The above-mentioned computer programs, plus programs to predict flexibility, location of transmembrane regions, Asn-linked glycosylation sites, and sites of signal sequence cleavage, are all contained in a package called PC Gene produced by IntelliGenetics (*APPENDIX 5*).

Selection of the Length of the Peptide

Generally, peptides with a length of 10 to 15 residues are used to make antipeptide sera that react with the native protein. Peptides with as few as 6 or as many as 35 amino acids have worked successfully; however, both extremes have disadvantages. Small peptides are more soluble and can produce very specific antisera, but the antibodies elicited by them are not as likely to react with the parent protein. Large peptides tend to be less soluble, more difficult to prepare synthetically, and are more likely to assume structures unrelated to the native protein. Part of the decision about peptide size will be determined by the individual peptide sequence, as some residues will adversely affect solubility, produce synthesis problems, or interfere with coupling.

Thus, in summary, a reasonable order of suggestions for choosing peptide sequences for making antipeptide sera would be:

1. If possible, use more than one peptide.
2. Use the C-terminal sequence (7 to 15 residues) if it is hydrophilic and if a suitable coupling group is available or can be added.
3. Use the N-terminal sequence (7 to 15 residues) if it is hydrophilic and if a suitable coupling group is available or can be added.
4. Use internal hydrophilic regions, perhaps using longer peptides (15 to 20 residues).

Modification of the Chosen Peptide

Other features of the peptide must be considered in order for it to mimic the native antigen as closely as possible. If the desired peptide sequence comes from an internal portion of the native protein, then the free N-terminal amino and C-terminal carboxyl groups (which are normally peptide-bonded to adjacent amino acid residues within the native protein) can be modified to more closely

Table 9.3.1 Hydrophobic/Hydrophilic Index of Amino Acids

Amino acid	Hydrophilicity value[a]	Hydropathy index[b]
Arginine (R)	3.0	−4.5
Aspartic acid (D)	3.0	−3.5
Glutamic acid(E)	3.0	−3.5
Lysine (K)	3.0	−3.9
Serine(S)	0.3	−0.8
Asparagine (N)	0.2	−3.5
Glutamine (Q)	0.2	−3.5
Glycine (G)	0.0	−0.4
Proline (P)	0.0	−1.6
Threonine (T)	−0.4	−0.7
Alanine (A)	−0.5	1.8
Histidine (H)	−0.5	−3.2
Cysteine (C)	−1.0	2.5
Methionine (M)	−1.3	1.9
Valine (V)	−1.5	4.2
Isoleucine (I)	−1.8	4.5
Leucine (L)	−1.8	3.8
Tyrosine (Y)	−2.3	−1.3
Phenylalanine (F)	−2.5	2.8
Tryptophan (W)	−3.4	−0.9

[a]Hopp and Woods (1981).
[b]Kyte and Doolittle (1982).

Table 9.3.2 Principal Carriers Used for Coupling Peptides[a]

Carrier[b]	M_r (kDa)	Number of groups/molecule			
		ε-NH_2	-SH	Phenol	Imidazole
BSA	67	59[c]	1	19	17
Ovalbumin	43	20	4	10	7
Myoglobin	17	19	0	3	12
Tetanus toxoid	150	106	10	81	14
KLH	>2000	6.9[d]	1.7[d]	7.0[d]	8.7[d]

[a]Adapted from Van Regenmortel et al. (1988).

[b]Abbreviations; BSA, bovine serum albumin; KLH, keyhole limpet hemocyanin.

[c]Only 30 to 35 of the 59 Lys residues of BSA are accessible.

[d]For KLH, the amino acid groups are expressed in grams of amino acid containing this functional group per 100 g.

mimic their native structure. The N-terminal amino group can be modified by acetylation of the peptide α-amino group during synthesis, and the C-terminal carboxyl group can be modified with a C-terminal amide during peptide synthesis. It is not certain that acetylation of the N terminus or formation of the C-terminal amide for peptides derived from internal sequences will really improve the chances of producing antisera reactive with the native protein. However, it has been demonstrated that these modifications will stabilize an α-helical conformation and may increase the solubility of the peptide.

SELECTION OF A CARRIER PROTEIN

A carrier protein should be a good immunogen and have a sufficient number of amino acid residues with reactive side-chains (see Table 9.3.2) for coupling to the synthetic peptide. KLH is commonly used because of its proven efficacy. However, because of its large size, KLH is more likely to precipitate during cross-linking, making the complex difficult to handle. Other proteins that have been used as carrier molecules include thyroglobulin, bovine serum albumin (BSA), and tetanus toxoid. BSA has the disadvantage that anti-carrier protein antibodies present in the anti-peptide serum will be a problem if the antiserum is used in the presence of fetal calf serum. Table 9.3.2 lists the most common protein carriers and their relevant properties.

SELECTION OF A COUPLING METHOD

In addition to the choice of peptide, a method for coupling the peptide to a protein carrier must also be selected. Most coupling methods rely on the presence of free amino (Lys), sulfhydyl (Cys), phenolic (Tyr), or carboxylic (Asp or Glu) groups. The chosen coupling method should link the peptide to the carrier via either the C- or N-terminal residue. Peptides corresponding to the amino terminus of proteins should be coupled through their carboxyl-terminal amino acid residue, whereas peptides corresponding to the carboxyl terminus of proteins should be coupled through their amino-terminal amino acid residue.

One coupling procedure that has proved to be particularly effective employs MBS as the coupling reagent. This procedure requires a free sulfhydryl group on the synthetic peptide and free amino groups on the carrier protein. Therefore, in order to use this method, it is usually necessary to add a Cys residue (during peptide synthesis) to the C or N terminus of the peptide. This will provide the sulfhydryl group for coupling to the carrier protein. Coupling a peptide derived from an N-terminal sequence to a carrier is accomplished with MBS via a Cys residue added to the C terminus of the peptide. This cross-links the peptide derived from the N-terminal sequence to the carrier molecule so that the N terminus is exposed as it would be in the native antigen. For a peptide derived from a C-terminal sequence, the Cys is placed on the N-terminus of the peptide for the same rationale. If an internal Cys residue is present in the chosen peptide, it can be used for MBS coupling, especially if the cysteine is part of a disulfide linkage in the native protein.

If coupling with MBS is not desirable for some reason (such as the presence of non-terminal Cys residues that are not disulfide-linked in the native protein), coupling can then be accomplished through a C- or N-terminal Tyr using *bis*-diazotized benzidine (BDB), through the C or N terminus with 1-ethyl-3-(3-dimethylaminopropyl)-carbodiimide (EDCI), or

Table 9.3.3 Principal Reagents Used for Peptide Protein Conjugation[a]

Coupling agent	Modified amino acid	
	Primary reaction	Secondary reaction
Glutaraldehyde	ε-NH_2Lys, α-NH_2, SH-Cys	Tyr, His
Bis-imido esters	α-NH_2, ε-NH_2Lys	Negligible
BDB	Tyr, SH-Cys, His, ε-NH_2Lys	Trp, Arg
Carbodiimides (EDCI)	α-NH_2, ε-NH_2Lys, α-COOH, Glu, Asp	Tyr, Cys
MBS	Cys-SH	Not observed

[a]Adapted from Van Regenmortel et al. (1988). Abbreviations: BDB, *bis*-diazotized benzidine; EDCI, 1-ethyl-3-(3-dimethylaminopropyl)-carbodiimide;MBS, *m*-maleimidobenzoyl-*N*-hydroxysuccinimide.

through the N-terminal α-amino group with glutaraldehyde. BDB coupling is not advisable if there are internal Tyr residues in the peptide. EDCI coupling is not advisable when internal Glu, Asp, or Lys residues are present. Glutaraldehyde coupling may not be appropriate if there are internal Lys residues in the peptide. Table 9.3.3 lists the principal reagents used for peptide-protein conjugation and the functional groups involved.

For example, if the carboxy-terminal sequence of a protein is

1 2 3 4 5 6 7 8 9 10
S Y G R N Q A E C Q—COOH

then coupling via MBS by adding a Cys residue to the N terminus may not be appropriate because of the Cys at position 9 (see Table 9.3.1 for single-letter codes). In this case, it may be preferable to couple the peptide via the N terminus using glutaraldehyde. However, if the Cys at position 9 is known to be part of a disulfide loop in the native protein, it may be better to couple with MBS through the naturally occurring Cys at position 9. Protocols for the coupling methods discussed here are presented in *UNIT 9.4*.

LITERATURE CITED

Chou, P.Y. and Fasman, G.D. 1974. Prediction of protein conformation. *Biochemistry* 13:222-245.

Corrigan, A.J. and Huang, P.C. 1982. A basic microcomputer program for plotting the secondary structure of proteins. *Comput. Programs Biomed.* 3:163-168.

Dyson, H.J., Cross, K.J., Houghten, R.A., Wilson, I.A., Wright, P.E., and Lerner, R.A. 1985. The immunodominant site of a synthetic immunogen has a conformational preference in water for a type-II reverse turn. *Nature (London)* 318:480-483.

Hopp, T.P. and Woods, K.R. 1981. Prediction of protein antigenic determinants from amino acid sequences. *Proc. Natl. Acad. Sci. U.S.A.* 78:3824-3828.

Hopp, T.P. and Woods, W.R. 1983. A computer program for predicting protein antigenic determinants. *Mol. Immunol.* 20:483-489.

Karplus, P.A. and Schulz, G.E. 1985. Prediction of chain flexibility in proteins: A tool for the selection of peptide antigens. *Naturwissenchaften* 72:212-213.

Kyte, J. and Doolittle, R.F. 1982. A simple method for displaying the hydropathic character of a protein. *J. Mol. Biol.* 157:105-132.

Van Regenmortel, M.H.V., Briand, J.P., Muller, S., and Plaué, S. 1988. Synthetic polypeptides as antigens. *In* Laboratory Techniques in Biochemistry and Molecular Biology, Vol. 19. (R.H. Burdon and P.H. von Knippenberg, eds.). Elsevier, Amsterdam and New York.

von Heijne, G. 1986. A new method for predicting signal sequence cleavage sites. *Nucl. Acids Res.* 14:4683-4690.

Westhof, E., Altschuh, D., Moras, D., Bloomer, A.C., Mondragon, A., Klug, A., and Van Regenmortel, M.H.V. 1984. Correlation between segmental mobility and the location of antigenic determinants in proteins. *Nature (London)* 311:123-126.

KEY REFERENCE

Van Regenmortel et al., 1988. See above.

Presents detailed discussion of selection of peptides for preparing antipeptide sera reactive with native proteins.

Contributed by W. Lee Maloy
Magainin Sciences
Plymouth Meeting, Pennsylvania

John E. Coligan
National Institute of Allergy and Infectious Diseases
Bethesda, Maryland